# Does gemination resist linguistic attrition?
# A study on Italian migrant speech in Melbourne Australia

Valentina De Iacovo[1], Paolo Mairano[2], John Hajek[3]

[1]Università di Torino, [2]Université de Lille, [3]The University of Melbourne
valentina.deiacovo@unito.it, paolo.mairano@univ-lille.fr, j.hajek@unimelb.edu.au

## ABSTRACT

In this study, we explore the possible effect of linguistic attrition on consonant gemination among speakers of Italian living in Australia. We analyse a corpus of spontaneous speech for a group of native Italian speakers who emigrated to Melbourne in Australia between the 1950s and 1970s. Given gemination is a fully lexicalised and stable phenomenon in Italian but not in English, we investigate whether, after many years of living in Australia, the phonological contrast between singleton and geminate consonants in Italian is preserved, or undergoes phonetic shift due to long-term contact with English. An analysis of acoustic durations reveals that all participants produced a significant difference in singleton and geminate consonant length, albeit with variable lengthening ratios. A post-hoc qualitative analysis seems to point to shorter geminate durations for speakers who have lived longer in Australia, but this will need to be supported by further research.

Keywords: gemination, L1 attrition, speech production, bilingualism, Italian.

## 1. THEORETICAL BACKGROUND

### 1.1 The Italian migrant community in Australia

The city of Melbourne has long been home to the largest Italian community in Australia as a result of mass migration from Italy that took place mainly between the 1950s and 1970s [1]. First generation migrants are today typically aged between 60 and 90 years old, and present a peculiar situation in terms of multilingualism/triglossia: they usually speak, as a result of simultaneous acquisition, (a) a regional dialect as spoken in their original village or hometown, (b) a regional variety of standard Italian supported by schooling in Italy before departure, as well as by interactions with other Italians in Australia, in addition to (c) an Australian-based variety of English which they have learnt later in life in order to integrate into Australian society. Subsequent generations mainly speak Australian English, and potentially also Italian, and, less frequently dialect, as heritage language(s) (HL) [2, 3]. The general picture described in the literature (e.g. [4]) reveals a complex and multifaceted linguistic situation, whereby the L1 Italian of these migrants can undergo different degrees of attrition, seen in a gradual decline in native language proficiency [5] according to the degree of influence of their later acquired L2 English, which also depends on social factors such as age of acquisition, level of education, family background, the use of English at work or for university studies. The linguistic studies carried out so far on Italians in Australia have often adopted a contrastive approach across generations [6] taking into account linguistic change and adaptation, with a particular focus on language shift [2] and code-switching [7]. While they are often far-reaching in scope, existing studies have noticeably paid little or no attention to the consequences of long-term language contact and shift on the phonetics and phonology of the Italian spoken by the first generation. The few exceptions to explore phonetic and/or phonological aspects of potential long-term attrition in L1 Italian outside Europe include [8], [9], [10] and [11]. The present contribution aims to add to this research by investigating potential phonetic and phonological attrition through the lens of a specific phonological feature of Italian: contrastive consonant gemination.

### 1.2 Linguistic attrition from a phonetic and phonological perspective

While linguistic attrition can impact on all areas of language proficiency and use, we focus our attention here on the phonetic and phonological aspects of it [12, 13, 14]. The term 'phonetic attrition' refers to L2-influenced phonetic change in an individual's L1 system over the long-term, most typically with respect to migrants living in an L2 context [12]. The related concept of 'phonological attrition' refers to a reduction in the phonetic robustness (erosion) and even loss of a contrast [8, 13] and is particularly relevant here given our focus on the impact of long-term L2 contact on the maintenance or otherwise of the consonant length contrast in Italian.

## 1.3 Consonant gemination in Italian vs English

Gemination consists of a temporal extension of a sound. It is considered as typologically marked, since it is phonologically contrastive in relatively few languages of the world. One of these languages is Italian, where gemination is distinctive within and across morphemes for 15 consonants [15], namely /p/, /b/, /t/, /d/, /k/, /g/, /tʃ/, /dʒ/, /f/, /v/, /s/, /m/, /n/, /l/, /r/; additionally, it is allophonic for 5 other consonants (/ts/, /dz/, /ʃ/, /ʎ/, /ɲ/), while the 3 remaining consonants (/z/, /j/, /w/) are always short. An illustrative minimal pair *cane* (dog) /ˈkane/ [ˈkaːne] and *canne* (canes) /ˈkanːe/ [ˈkanːe] shows the gemination contrast. The duration of the geminate consonant is greater than that of singleton by a ratio of approx. 2:1 [16], with some variation depending on the category of sounds, [17, 18]; large-scale studies on running and spontaneous speech found a smaller ratio of approx. 1.7:1 [19]. Gemination also affects the preceding vowel, which is shortened before a geminate consonant (but only in nuclear position, cf. [19] and [20]). Additionally, gemination can also occur in standard Italian as an external sandhi phenomenon (referred to as *raddoppiamento fonosintattico*), but this phenomenon is not present in all varieties of Italian and will not be analysed in our study.

English, on the contrary, does not have a lexicalized long-short consonant contrast within morphemes, although in more careful styles there may be so-called *fake* gemination due to the concatenation of two morphemes ending and starting with the same sound, such as in *unnatural*. These fake geminates, not characteristic of all speakers, are also typically reduced in spontaneous speech (see [21] for detailed discussion and overview).

## 1.4 Gemination and attrition

Since gemination is typologically marked, and as such is restricted or absent in many languages, it is not unreasonable to hypothesise that it may undergo some degree of erosion for Italian migrants who have lived for many years in Australia. In fact, the erosion of the phonological contrast between singleton and geminate consonants in migrants and heritage speakers, as a result of contact with English, has been investigated in a series of production studies for L1 Farsi and L1 Arabic speakers living in Canada [22 and 23 respectively] and in the US [24]: their results point to evidence of attrition for this phonological contrast and a gradual shift towards English, whereby the acoustic duration of L1 geminate consonants is on average still longer than for singletons, but shorter than for geminate consonants produced by monolingual Farsi and Arabic speakers. Interestingly,

generation was found to be a predictor of the degree of erosion, showing that it increased in successive generations (for example, [22] considered generations 1, 1.5 and 2). On the other hand, other studies have not uncovered evidence of geminate attrition in migrant speech: [25] found no differences in geminate production for HL speakers of Italian in Germany, while [26] did not find changes in the perception of geminates for L1 Hungarian speakers living in Germany and the US. [8] explored (de-)gemination in Italian migrants and heritage speakers in the US, and more particularly the perception of geminates by members of the Italian community from Lucca now living in San Francisco, CA. The authors found evidence of a progressive impairment in the perception of the length feature across generations, as well as differences between first generation immigrants and Lucchese speakers still living in the area around Lucca. No study has yet examined the production of geminate consonants in L1 Italian migrants or heritage speakers.

## 1.5 Aim of this study

In this contribution we present the results of a study aimed at exploring potential erosion patterns of gemination for L1 Italian migrants in Australia. As noted above, the existing studies focusing on the erosion of L1 gemination are scarce and report conflicting results: evidence of erosion is found for L1 Farsi [22], L1 Arabic ([23], [24]), L1 Italian [8] speakers, but not for Hungarian [26] speakers. For this preliminary study, we focus on whether first generation L1 Italian migrants in Australia produce geminates at all, and if so, whether geminate durations are comparable to those produced by L1 Italian speakers living in Italy.

## 2. DATA AND METHODS

### 2.1 Corpus

For this study 10 first-generation bilingual Italian-Australian speakers living in Melbourne (Australia) were recruited. The speakers (6 females and 4 males) are aged between 61 and 91 and have lived in Australia around 43 and 67 years. They were born in different parts of Italy and for the most part grew up there. They moved, at different ages, to Australia, where they learned English. Only two participants arrived relatively early in Australia (ages 6 and 7 respectively) and belong to what is sometimes referred to as generation 1.5. Upon arrival in Australia, four of them pursued their education/studies in English, while all participants worked in a mixed Italian-English environment, maintaining contact with members of the Italian

community and with their families. All participants have continued to speak Italian, as well as English, on a regular basis during their time in Australia. The speech material used for analysis is now contained within the so-called *SEIM* corpus, consisting of semi-spontaneous audio-visual interviews recorded in Italian by the first author with Italian bilinguals and heritage speakers with the aim of creating an oral archive for speech analysis. The interviews were conducted by following a set of questions addressing specific topics relating to their experience as migrants (arrival in Australia, lifestyle, use of Italian). Most speakers were interviewed individually by the first author, while three of them (speakers BE, AN, MA) were interviewed in each case with a partner – but their partner's speech was not analysed. The recordings were made using the microphone of an *IPad mini* at a sampling rate of 48 kHz. They have variable durations, mostly around 30-40 minutes. For the present analysis we considered the first 10 minutes of speech for each speaker.

**2.2 Acoustic measurements**

The audio files were extracted in wav format and resampled at 16kHz. The sound material was first transcribed orthographically manually, then phonetized and forced aligned using WebMaus [27], resulting in *textgrids*. Segmentation was subsequently checked manually in *Praat* [28]. The duration values of the target sounds were then extracted for all 15 geminating consonants (/tʃ, dʒ, p, t, k, b, d, g, m, n, f, s, v, r, l/), in word-medial contexts where gemination may potentially occur according to phonotactic restrictions of Italian, i.e. preceded by a vowel, and followed by a vowel, liquid or approximant. Consonants with few geminate occurrences were eliminated from the analysis (namely affricates, fricatives and voiced plosives, thereby leaving voiceless plosives, nasals and liquids). Due to limitations of space, we report only on the total number of tokens analysed per speaker (table 1) without further distinction by manner or other segment type.

| Speaker | Singletons | Geminates |
|---------|-----------|-----------|
| AD | 442 | 86 |
| AC | 503 | 112 |
| AP | 298 | 49 |
| AM | 387 | 103 |
| BM | 463 | 121 |
| CP | 441 | 138 |
| CM | 365 | 74 |
| CT | 549 | 120 |
| GB | 486 | 105 |
| MP | 547 | 147 |

**Table 1**: Number of tokens analysed per speaker.

# 3. RESULTS

The data were then imported into *R* [29], and analysed with linear mixed-effects models with the *lme4* [30] library, *lmeRtest* [31] was used to obtain *p* values. The structure of random slopes was kept maximal.

Inspection of the data (figure 1) immediately illustrates the presence of geminate consonants in the speech of our speakers. This was also confirmed by a first generic model testing durational differences between singletons and geminates in our data, including speaker, sound and word as random factors: *duration ~ isGeminate + (isGeminate/speaker) + (isGeminate/sound) + (1/word)*. Sound was included as a random effect to account to the fact that different consonants have intrinsically different durations. This model clearly illustrated that geminates are longer than singletons by 31 ms on average ($SE = .004$, $t = 8.097$, $p < .001$) for our speakers. A second generic model with the same structure tested the duration of the preceding vowel, but in this case no significant effect was found. This is not surprising and simply reflects results on running speech by previous studies (e.g. [19]).
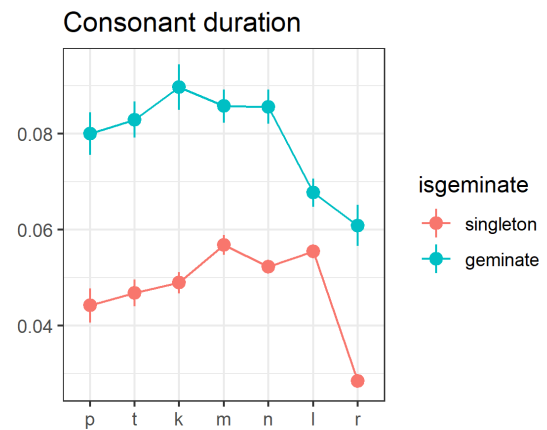


**Figure 1**: Mean consonant durations averaged across speakers.

Subsequently, we built a third model to investigate whether all our speakers produce geminate consonants, and with which ratios. With this in mind, we transformed speaker into a fixed (rather than random) variable, in order to be able to perform post-hoc analyses on each speaker, and ran the following model on consonant durations: *duration ~ isGeminate * speaker + (isGeminate/sound) + (1/word)*. An Anova of the model confirmed that the main effect of geminate status on consonant durations was highly significant ($F = 88.28$, $p < .001$), but also that consonant durations varied by speaker ($F = 34.14$, $p < .001$), and that there was an interaction between geminate status and speaker ($F = 9.43$, $p < .001$). A post-hoc analysis with Holm correction for multiple

comparisons revealed that the singleton-geminate difference in duration was significant for every single participant (all *p* values < .001). The predicted values for singletons and geminates for each speaker were then extracted from the model and are plotted in figure 2: these values illustrate the existence of a clear-cut distinction between geminates and singletons for every speaker.
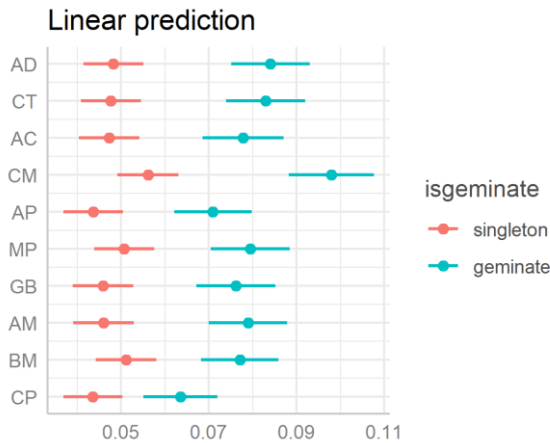
**Figure 2**: Linear prediction of geminate and singleton durations in sec. for every speaker.
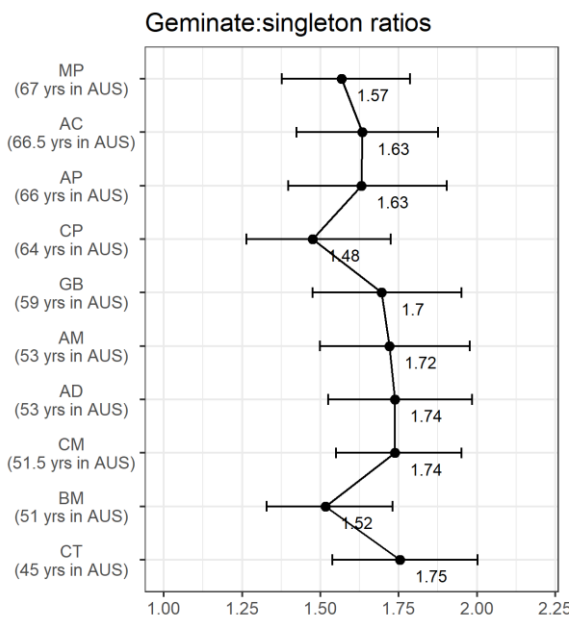
**Figure 3**: By-speaker geminate:singleton ratios computed on linear predictions for consonant durations.

Since the plot in figure 2 shows that the magnitude of lengthening is not the same for all participants, we used these values to compute by-speaker geminate:singleton ratios. These are shown in figure 3, where speakers are sorted by length of residence in Australia. This plot shows that most speakers have a geminate:singleton ratio between 1.6 and 1.7, which is comparable to what was found by [19] for running speech involving native speakers of Italian in Italy.

Only two speakers stand out as having smaller ratios (1.52:1 and 1.48:1), namely CP and BM. This may be an indication of erosion for these two speakers. Interestingly, the two 1.5 generation speakers (AC and AD) do not stand out; instead, they behave in a similar fashion to first generation speakers in terms of gemination.

## 4. FINAL DISCUSSION

Our analysis of singleton and geminate consonant durations of L1 Italian migrants in Australia revealed little if any attrition: (1) our speakers maintain the geminate contrast, and (2) they generally do so with a magnitude of differentiation comparable to native speakers in Italy, i.e. with geminate:singleton ratios of 1.6:1 and 1.7:1 (similar to results obtained for running speech by [19]). This also includes our two 1.5 generation speakers, who do not show any patterns of gemination specific only to them. These last results appear to be consistent with [25], who report that geminates produced by their German-dominant Italian HL speakers were reliably longer than their singletons. That said, we do observe smaller ratios of gemination for two other speakers in our sample, which may be due to linguistic attrition; this may have various causes, and may depend on specific linguistic behaviours or on different permeability to linguistic erosion [12].

Finally, we also note that speakers who have lived in Australia for less than 60 years tend to have ratios of 1.7 or higher (with the exception of BM), while speakers with more than 60 years of residence show ratios of less than 1.7. This may be the indication of a trend to decrease the magnitude of lengthening for geminate consonants after many years of residence in Australia. The correlation between the number of years in Australia and geminate:singleton ratios is however non-significant ($R = -.54$, $p = .107$) and this trend could be an artefact due to the small number of participants available. Future larger-scale and more controlled studies may give us a deeper insight into this issue. For now, we are able to conclude that L1 Italian speakers in Australia produce geminate consonants with patterns that are comparable to those of native speakers in Italy, with lengthening ratios of 1.6-1.7. If erosion does apply to geminate consonants at all, then it may do so only at very late stages, or only for a minority of speakers.

# 6. REFERENCES

[1] Clyne, M., Hajek, J., Kipp, S. 2008. Tale of two multilingual cities in a multilingual continent. *People and Place* 16, 1–9.

[2] Bettoni, C. 1989. Language Loss Among Italians in Australia. *International Journal of Applied Linguistics* 83(1), 37–50.

[3] Clyne, M. 1987. Bilingualism and community languages: The educational consequences. *L'italiano in Australia* 31(1-2), 57–71.

[4] Bettoni, C. 1989. Maintaining and developing Italian in Australia. *Vox, The Journal of the Australian Advisory Council on Languages and Multicultural Education* 3, 48–53.

[5] Schmid, M. S. 2009. On L1 attrition and the linguistic system. *EuroSLA yearbook* 9(1), 212–244.

[6] Bettoni, C., Gibbons, J. 1988. Linguistic purism and language shift: a guise-voice study of the Italian Community in Sydney. *International Journal of the Sociology of Language* 72, 15–35.

[7] Rubino, A. 2004. Trilingual women as language mediators in the family: a Sicilian-Australian case study. *Australian Review of Applied Linguistics* 18, 25–49.

[8] Celata, C., Cancila, J. 2010. Phonological attrition and the perception of geminate consonants in the Lucchese community of San Francisco (CA). *International Journal of Bilingualism* 14, 1–25.

[9] Tordini, O., Galatà, V., Avesani, C., Vayra, M. 2018. Sound maintenance and change: Exploring inter-language phonetic influence in first-generation Italo-Australian immigrants. Il parlato nel contesto naturale: Modelli e metodi per lo studio del parlato in reali condizioni comunicative—Speech in the natural context. Models and methods for the analysis of speech under real communicative conditions. *Collana Studi AISV* 4, 77–98.

[10] Nodari, R., Celata, C., Nagy, N. 2019. Socio-indexical phonetic features in the heritage language context: Voiceless stop aspiration in the Calabrian community in Toronto. *Journal of Phonetics* 73, 91–112.

[11] Frontera, M. 2020. Calabrian migrants in Argentina: assessing first language phonetic attrition. *Proc. 16th AISV Conference* Rende (CS), 73–94.

[12] De Leeuw, E. 2019. Phonetic attrition. In: Schmid, M. S., Köpke, B. (eds), *The Oxford handbook of language attrition*. Oxford University Press, 204–217.

[13] Celata, C. 2019. Phonological attrition. In: Schmid, M. S., Köpke, B. (eds), *The Oxford handbook of language attrition*. Oxford University Press, 218–227.

[14] Chang, C. B. 2019. Phonetic drift. In: Schmid, M. S., Köpke, B. (eds), *The Oxford Handbook of Language Attrition*. Oxford University Press, 191–203.

[15] Loporcaro, M. 1996. On the analysis of geminate in Standard Italian and Italian dialects. In: Hurch, B., Rhodes, R. (eds), *Natural phonology: The state of the art.* Mouton de Gruyter, 153–187.

[16] Esposito, A., Di Benedetto, M. G. 1999. Acoustical and perceptual study of gemination in Italian stops. *The Journal of the Acoustical Society of America* 106(4), 2051–2062.

[17] Di Benedetto, M. G., De Nardis, L. 2021. Consonant gemination in Italian: The nasal and liquid case. *Speech Communication* 133, 62–80.

[18] Di Benedetto, M. G., De Nardis, L. 2021. Consonant gemination in Italian: The affricate and fricative case. *Speech Communication* 134, 86–108.

[19] Mairano, P., De Iacovo, V. 2020. Gemination in Northern versus Central and Southern Varieties of Italian: A Corpus-based Investigation. *Language and Speech* 63(3), 608–634.

[20] Bertinetto, P. M., Dell'Aglio, M., Agonigi, M. 2008. Quali fattori influenzano maggiormente la durata vocalica e consonantica in italiano? Un'indagine mediante l'algoritmo di decisione C&RT. *La comunicazione parlata*, 1000–1026.

[21] Kaye, A.S. 2005. Gemination in English. *English Today* 21(2), 43–55.

[22] Rafat, Y., Mohaghegh, M., Stevenson, R.A. 2017. Geminate attrition across three generations of Farsi-English bilinguals living in Canada: An acoustic study. *Ilha do Desterro Investigating Second Language Speech Learning* 71(3), 151–168.

[23] Alkhudidi, A., Rafat, Y., Stevenson, R. 2018. Geminate Attrition in the Speech of Arabic- English Bilinguals Living in Canada. *Heritage Language Journal* 17(1), 1–37.

[24] Hanini, R., Alkhudidi, A., Rafat, Y., Spinu, L. 2019. Geminate Attrition in the Speech of Arabic-English Bilinguals Living in the United States. *178th meeting of Acoustical Society of America* San Diego, CA.

[25] Einfeldt, M., van de Weijer, J., Kupisch, T. 2018. The production of geminates in Italian-dominant bilinguals and heritage speakers of Italian. *Language, Interaction and Acquisition* 10(2), 177–203.

[26] Neuberger, T. 2022. Effects of L2 exposure on the perception of the singleton/geminate contrast in Hungarian. *Alkalmazott Nyelvtudomány* 22(1), 191–207.

[27] Kisler, T., Uwe, R., Schiel, F. 2017: Multilingual processing of speech via web services, *Computer Speech & Language* 45, 326–347.

[28] Boersma, P., Weenink, D. 2022. Praat: doing phonetics by computer [Computer program]. Version 6.3.03, retrieved 17 December 2022 from http://www.praat.org/)

[29] R Core Team (2022) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

[30] Bates, D., Maechler, M., Bolker, B., Walker, S. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67, 1–48.

[31] Kuznetsova, A., Brockhoff, P.B., Christensen, R.H.B. 2017. lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software* 82, 1–26.